

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Studi tentang temuan terhadap confusion matrix, bersama dengan hasil tes dan penilaian dalam membandingkan kedua algoritma, memungkinkan untuk mengambil kesimpulan perbandingan antara OLR dan NN:

1. Berdasarkan data kualitas udara Jakarta, penelitian ini menunjukkan bahwa enam metrik utama (PM10, PM2.5, SO₂, CO, O₃, dan NO₂) dapat digunakan untuk menentukan kategori kualitas udara yang baik, sedang, tidak sehat, dan berbahaya. Kinerja model yang dibangun ditingkatkan oleh prosedur persiapan data seperti pembersihan data, penanganan nilai yang hilang, normalisasi, dan penyeimbangan kelas. Ini menunjukkan bahwa kualitas hasil analisis sangat dipengaruhi oleh kualitas dataset yang digunakan.
2. Hasil pengujian menggunakan algoritma Ordinal Logistic Regression (OLR) menunjukkan akurasi yang cukup baik, antara 91 hingga 92 persen. OLR dapat memprediksi dengan cukup baik, terutama untuk beberapa kategori, tetapi memiliki kelemahan dalam menyebarkan prediksi secara merata. Model ini hanya 'kuat' di satu kelas dan kurang stabil di kelas lainnya, menurut matriks kebingungan. Kekurangan OLR dalam mengelola data dengan hubungan yang rumit dan non-linier adalah penyebab masalah ini.
3. Dengan tingkat akurasi hingga 93%, hasil tes yang menggunakan algoritma Jaringan Saraf (NN) menunjukkan kinerja yang lebih baik dibandingkan OLR. NN dapat mendistribusikan hasil prediksi secara lebih merata di semua kategori kualitas udara selain memiliki akurasi yang lebih tinggi.

Ketika hiperparameter diubah menggunakan Grid Search, model dapat menentukan kombinasi konfigurasi yang optimal, meningkatkan hasil tes dan menyoroti manfaat dari NN.

4. Dari perbandingan kedua algoritma, jelas bahwa Jaringan Saraf lebih unggul dibandingkan Regresi Logistik Ordinal dalam klasifikasi kualitas udara di Jakarta. Ini disebabkan oleh kemampuan Jaringan Saraf untuk mengidentifikasi pola rumit dan korelasi non-linear antara variabel yang mempengaruhi polutan. Akibatnya, perkiraan yang dihasilkan lebih akurat dan dapat diandalkan. Oleh karena itu, jaringan saraf lebih cocok untuk penelitian dan aplikasi yang memerlukan klasifikasi kualitas udara yang lebih tepat.

5.2 Saran

1. Hanya data kualitas udara dari tahun 2021–2024 yang digunakan dalam analisis ini, dan dibatasi pada enam parameter polutan. Data tersebut sebaiknya diperluas untuk mencakup periode waktu yang lebih lama dan menggabungkan faktor meteorologi seperti suhu, kelembapan, dan kecepatan angin jika penelitian lebih lanjut dilakukan. Sebagai hasilnya, model tersebut akan dapat memberikan prediksi yang lebih komprehensif.
2. Selain OLR dan jaringan saraf tiruan, ada banyak teknik pembelajaran mesin lainnya yang dapat diterapkan. Ini termasuk model penangkapan pola deret waktu yang didasarkan pada pembelajaran dalam, seperti Long Short-Term Memory (LSTM), Random Forest, dan Gradient Boosting. Temuan penelitian akan lebih komprehensif dan mampu mengusulkan model mana yang paling sesuai untuk kondisi kualitas udara di Jakarta jika ada lebih banyak perbandingan algoritma.
3. Model yang dibuat dalam studi ini saat ini terbatas pada data historis. Oleh karena itu, akan lebih baik jika model tersebut diterapkan pada sistem yang menggunakan teknologi Internet of Things (IoT) untuk memantau kualitas

udara secara real-time. Dengan demikian, publik dapat langsung menggunakan hasil klasifikasi sebagai sistem peringatan.

Diharapkan bahwa studi ini tidak hanya akan memberikan manfaat bagi akademisi tetapi juga membantu pemerintah, khususnya Badan Lingkungan Hidup DKI Jakarta, dalam mengembangkan langkah-langkah pengendalian polusi udara. Kampanye untuk meningkatkan kesadaran mengenai risiko yang ditimbulkan oleh polusi udara dan pentingnya menjaga lingkungan juga harus digunakan untuk melibatkan masyarakat.

